《Statistics for Large-Dimensional Data》 Glossary

Bootstrap resampling: 【ブートストラップ・リサンプリング】 A data-based simulation method for statistical inference which can be used to study the variability of estimated characteristics of the probability distribution of a set of observations and to provide confidence intervals for parameters in situations where these are difficult or impossible to derive in the usual way. (The use of the term “bootstrap” derives from the phrase ‘to pull oneself up by one’s bootstraps’.)

Kernel: 【カーネル】 A function of two variables, written for example by $k(x, y)$.

Positive definite — (real-valued): 【正定値—】 A real-valued function $k(x, y)$ of two variables on a set Ω such that (i) it is symmetric $k(x, y) = k(y, x)$, and (ii) for any number of points $x_1, ..., x_n$ on Ω the symmetric matrix $(k(x_i, x_j))_{i,j}$ is nonnegative definite. It is known that a positive definite kernel is associated with a Hilbert space consisting of functions on Ω such that the kernel k satisfies the reproducing property, that is, $f(x) = \langle f, k(\cdot.x) \rangle$ for any f in the Hilbert space and x in Ω, where $\langle , \rangle$ denotes the inner product. Such a Hilbert space is called a reproducing kernel Hilbert space associated with k.

Machine learning: 【機械学習】 A branch of artificial intelligence. It is a scientific discipline concerned with the design and development of algorithms that allow computers to evolve behaviors based on empirical data, such as from sensor data or databases.

Regression: 【回帰】 A statistical method that attempts to determine the strength of the relationship between one dependent variable (usually denoted by Y) and one or more changing variables X (known as independent variables). More specifically, regression analysis helps one understand how the typical value of the dependent variable changes when any one of the independent variables is varied, while the other independent variables are held fixed. Most commonly, regression analysis estimates the conditional expectation of the dependent variable given the independent variables $\mathrm{E}[Y|X]$ — that is, the average value of the dependent variable when the independent variables are held fixed.